

# Raster Images Procedures

Version 1.107

<b>Created date:</b>	26 January 2012
<b>Last updated:</b>	25 August 2017
<b>Review Due:</b>	25 August 2018
<b>Authors:</b>	Jen Mitcham, Stewart Waller, Jon Bateman, Kieron Niven, Michael Charno, Tim Evans, Jenny O'Brien and Ray Moore
<b>Maintained by:</b>	Digital Archivists
<b>Previous version:</b>	Live

## 1 Purpose of this document

This page is designed to document the current ADS procedures for production of dissemination and preservation copies of raster images. It contains a list of current dissemination/preservation formats and how to migrate files to these formats.

For more information on this data type, please refer to the *Guides to Good Practice for Raster Images*.<sup>1</sup>

**Please note for geo-referenced rasters please consult the Data Procedures for GIS.**

---

<sup>1</sup> [http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg\\_Toc](http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_Toc)

## 2 Formats

Offered format	Accepted	Preservation	Presentation	Notes
Uncompressed Baseline TIFF v.6 <b>.tif</b>	YES	Uncompressed Baseline TIFF v.6 <b>.tif</b>	Portable Network Graphics <b>.png</b> (lossless compression @ c.5) or Joint Photographic Expert Group <b>.jpg</b>	If EXIF & IPTC metadata is significant then this should be preserved during conversion, or if regarded as significant by the archivist and following negotiation with the depositor.
Portable Network Graphics <b>.png</b>	YES	Uncompressed Baseline TIFF v.6 <b>.tif</b>	Portable Network Graphics <b>.png</b>	a/a
Joint Photographic Expert Group <b>.jpg / .jpeg</b>	YES	Uncompressed Baseline TIFF v.6 <b>.tif</b>	Joint Photographic Expert Group <b>.jpg / .jpeg</b>	a/a
Graphics Interchange Format (Compuserve) <b>.gif</b>	YES	Uncompressed Baseline TIFF v.6 <b>.tif</b>	Portable Network Graphics <b>.png</b>	a/a
Bit-Mapped Graphics Format (Microsoft) <b>.bmp</b>	YES	Uncompressed Baseline TIFF v.6 <b>.tif</b>	Portable Network Graphics <b>.png</b>	a/a
PhotoCD <b>.pcd</b>	NO	N/A	N/A	We don't take these anymore
Photoshop (Adobe) <b>.psd</b>	NO	N/A	N/A	We don't take these anymore
CorelPaint <b>.cpt</b>	NO	N/A	N/A	We don't take these anymore
Adobe Digital Negative <b>.dng</b>	YES	Adobe Digital Negative <b>.dng</b> and Uncompressed Baseline TIFF v.6 <b>.tif</b>	Adobe Digital Negative <b>.dng</b> and Joint Photographic Expert Group <b>.jpg</b>	If EXIF & IPTC metadata is significant then this should be preserved during conversion, or if regarded as significant

				by the archivist and following negotiation with the depositor.
JPEG2000 <b>.jp2 / .jpx</b>	YES	Uncompressed Baseline TIFF v.6 <b>.tif</b>	JPEG2000 <b>.jp2 / .jpx</b>	a/a
PDF	YES, but only if not available in original format	Uncompressed Baseline TIFF v.6 <b>.tif</b> or Portable Document Format <b>.pdf</b>	Joint Photographic Expert Group <b>.jpg / .jpeg</b> or Portable Document Format <b>.pdf</b>	For where drawings, photos, scans etc. have been deposited as PDF files but would have a data type of image. Ideally these should be deposited in their original format (due to the loss of information resulting from the creation of the PDF. Preservation format is at the discretion of the archivist.
RAW	NO	N/A	N/A	

Please see the *Guides to Good Practice* page for more detailed discussion of these, and other, formats.<sup>2</sup>

### 3 Documentation / Metadata

The following documentation is required for Raster Images.

Element	Description
Filename	
Caption	
Subject keyword	Where appropriate these should be terms used in recognised thesauri (e.g. FISH, <sup>3</sup> Heritage Data: LOD, <sup>4</sup> LCSH, <sup>5</sup> etc.
Period Term	Where appropriate these should be terms used in recognised thesauri (e.g. FISH, <sup>6</sup> Heritage Data: LOD, <sup>7</sup> LCSH, <sup>8</sup> etc.

<sup>2</sup> [http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg\\_2](http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_2)

<sup>3</sup> <http://thesaurus.historicengland.org.uk/frequentuser.htm>

<sup>4</sup> <http://heritagedata.org/live/getAllSchemes.php>

<sup>5</sup> <http://id.loc.gov/authorities/subjects.html>

<sup>6</sup> <http://thesaurus.historicengland.org.uk/frequentuser.htm>

<sup>7</sup> <http://heritagedata.org/live/getAllSchemes.php>

<sup>8</sup> <http://id.loc.gov/authorities/subjects.html>

Period date	
Creator	
Copyright Holder	
Location	
Grid Reference	
Creation date	
Software	
Software version	

The *Guides to Good Practice* also recommend the following documentation is supplied for each file, but these are not currently requested but could be created by the archivist.<sup>9</sup>

Element	Description
File Format and Version	
File Size	Size of the file in bytes.
Spatial Resolution	The resolution (ppi).
Dimensions	Dimensions in pixels.
Colour Space	The colour space.
Bit Depth	e.g. 24bit or 8bit.
Capture Device	Details of camera or scanner.

## 4 Accessioning checks

- Do we have the necessary documentation (see below)
- Images are suitable for deposition i.e. they do not contain inappropriate content (e.g. children, people, etc). Copyright permissions should be assumed, but if questionable then this should be checked with the depositor.
- TIF files are uncompressed.
- Embedded metadata (e.g. EXIF, and other types) in file. **See section below**, if regarded as significant by the archivist and following negotiation with the depositor.
- It is quite common for raster content to be deposited as a single PDF file (or multiple images in a PDF file). In this case we should ask for the original (i.e. non-PDF file. If this is not possible, then we will have to proceed on a "best-efforts" basis
- Geo-rectified raster images - guidance on these is provided in the GIS procedures document

<sup>9</sup> [http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg\\_Toc](http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_Toc)

## Significant properties

From the *Guides to Good Practice*:

*The significant properties of raster images are discussed in detail in the InSPECT Significant Properties Testing Report on Raster Images (Montague 2009)...*

- **Image Size and Resolution** - conversions should ensure that the original resolution and image size remains the same in the preservation file format. In addition it is important that, when converting files to a new format, lossy compression is not applied to the image.
- **Bit depth and Colour space** - converted files should ensure that the bit depth and colour space of the original image are supported in preservation formats and that images are not degraded when converted.

*Although these properties are components of all image formats it is important to ensure that these properties remain the same/retain the same values when converting files to archival formats.<sup>10</sup>*

Embedded metadata such as EXIF and IPTC data, where present, also remains a significant property (see below).

## Embedded Metadata

From the *Guides*:

*... embedded metadata such as EXIF and IPTC can also be seen in certain cases as a significant property of an image and, where relevant, should be preserved with the file or exported to a separate plain or delimited text or XML file to be stored alongside the image. Although it is possible to preserve JPEG EXIF within the TIFF tag structure it is better held in a separate file, avoiding the risk of loss or corruption during later migration and making the metadata more easily accessible. Extraction of EXIF fields is relatively straightforward, with a number of free tools available.<sup>11</sup>*

As stated above, in an ideal world we should ask for such metadata to be supplied separately as XML or TXT. If we **do receive this** then this data should be preserved and disseminated with the relevant files (see below for notes on storage).

However, it is most likely that we will receive digital images - either taken with a digital camera or scanned - with embedded metadata but not in an additional format.

The process of embedding metadata within images, such as geo-tagging TIF files or producing EXIF information via a digital cameras, has become common practice. Care must be taken when converting or reproducing source images that this meta data be retained.

Work on the G2GP has shown that embedded metadata is often created automatically (usually technical information and things like dates). So, if it is noted that a file has embedded metadata, that is not also supplied as XML / TXT, **a decision should be made about the value of this information**. Generally, if embedded metadata is seen to be valuable, it should be preserved (see notes on storage below). As noted in the *Guides*, there are numerous free online tools to extract this information, such as EMET.<sup>12</sup>

---

<sup>10</sup> [http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg\\_3](http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_3)

<sup>11</sup> [http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg\\_3](http://guides.archaeologydataservice.ac.uk/g2gp/RasterImg_3)

<sup>12</sup> <http://www.artstor.org/form/emet>

When disseminating files, it is best practice to try and include the metadata in the dissemination format, thus allowing for more advanced reuse of the file. JPG keeps this information, PNG generally does not. Therefore choose your dissemination strategy appropriately.

## 5 How to convert files

Files should be converted according to the strategy chosen by the Digital Archivist.

### Extracting Embedded Metadata

If the Digital Archivist has decided that embedded metadata is a significant property of a file, then this should be extracted and saved in a preservation suitable format.

By far the best method of extracting this data is using *Adobe Bridge*:<sup>13</sup>

1. Download the file *BarredRock CSV Extract.jsx* attached to this page and save it to the Scripts folder (mine is at *Applications => Adobe => Bridge CS3 => Startup Scripts*)
2. Start Adobe Bridge
3. Navigate to the folder of images you want to extract metadata from, and select all images (so they're highlighted)
4. From the toolbar at the top of the screen select "Scripts => Export Metadata .."
5. You should be presented with a screen that allows you to export metadata by category (e.g. EXIF) and select the fields you want to include
  - Make sure 'Include filename' and 'Include Column names' are selected.
  - If IPTC metadata is present, Select the Schema "IPTC Core" and include all fields that need to be preserved
    - Click OK, a csv will be created in the same directory as the images - **follow procedures for file naming and storage**
  - If EXIF metadata is present, Select the Schema "EXIF" and include all fields that need to be preserved
    - Click OK, a csv will be created in the same directory as the images - **follow procedures for file naming and storage**

### Thumbnails

All images for download should have a thumbnail (for raster images, these should be jpgs). Thumbnail images should be a maximum width OR maximum height of **125px** - so that landscape and portrait images have the same surface area. Thumbnails should be renamed with the prefix "**thumb**" and store in the web folder under */images/thumbs*. Please note, these images **are not** part of the AIP and should never be stored with the actual data.

### Preview Images

In **ALL** cases, for example for large (bytes and pixels wise) a preview image will need to be created. The interface should present this as a "preview" via highslide, the user can then decide whether to download the full file. Preview images should be a maximum width OR maximum height of **750px** - so that landscape and portrait images have the same surface area. These should be stored in the web folder under */images/preview/*. Please note, these images **are not** part of the AIP and should never be stored with the actual data.

For more information/guidance on interface design please see the ADS Archive Interface page<sup>14</sup> and the Highslide page.<sup>15</sup>

---

<sup>13</sup> <https://www.adobe.com/uk/products/bridge.html>

## File naming

Where possible files should retain the same name as the original.

original\_name.jpg

Extracted metadata (created in Adobe Bridge)<sup>16</sup> will have to be named consistently, for example:

original\_name\_exif\_meta.csv

Then placed in prescribed location (see below).

## 6 Post-migration checking

We should ensure data consistency by undertaking a check of files post migration. The number of files to be checked is at the discretion of the Digital Archivist. Things to check include:

1. Bit depth and colour space of the original image are retained
2. Original resolution and image size remains the same - images are not degraded
3. Embedded metadata retained (if appropriate)

XnView<sup>17</sup> or Adobe Bridge<sup>18</sup> are good tools to use for this checking. You can view the properties of a file (Colour Model, Pixels per Inch etc), in the file properties tab of XnView or the Metadata "frame" in Adobe Bridge.

## Storage

Data should be stored as described in the ADS Repository Operations.<sup>19</sup> Any directory structure from the SIP should be retained in the AIP. In some cases editing/restructuring may be required, any restructuring must be recorded in the Process table in the CMS. Raster images will often be used to record metadata, usually an entity relationship model for a database.

Otherwise, store data in one of the following directory structure:

```
/preservation
  /{original_structure}
    myimage.tif
    myimage2.tif

/dissemination
  /{original_structure}
    myimage.jpg
    myimage2.png
```

---

<sup>14</sup> Available from the ADS internal wiki.

<sup>15</sup> Available from the ADS internal wiki.

<sup>16</sup> <https://www.adobe.com/uk/products/bridge.html>

<sup>17</sup> <https://www.xnview.com/en/>

<sup>18</sup> <https://www.adobe.com/uk/products/bridge.html>

<sup>19</sup> <http://archaeologydataservice.ac.uk/advice/RepositoryOperations.xhtml>

## Storing metadata

File metadata should be stored in an appropriate format with the preservation/dissemination files in a "documentation" folder. EXIF/ IPTC metadata For preservation purposes, any 'embedded' metadata should be stored separately in a "documentation" sub-folder as **.csv** (or **.xml** / **.txt**), for example:

```
/preservation
  /{original_structure}
    myimage.tif
  /documentation
    myimage_exif_meta.csv
    myimage_metadata.xlsx

/dissemination
  /{original_structure}
    myimage.jpg
  /documentation
    myimage_metadata.xlsx
```

## Raster Image metadata import

All deposits using the existing raster metadata template, available from the *Guidelines for Depositors*<sup>20</sup>, can use the import function within the CMS.<sup>21</sup>

---

<sup>20</sup> <http://archaeologydataservice.ac.uk/advice/FilelevelMetadata.xhtml#Raster%20Images>

<sup>21</sup> Available from the ADS internal wiki.